



## **Complete genome sequence of "Thiodictyon syntrophicum" sp. nov. strain Cad16T, a photolithoautotrophic purple sulfur bacterium isolated from the alpine meromictic Lake Cadagno**

Luedin, Samuel M.; Pothier, Joël F.; Danza, Francesco; Storelli, Nicola; Frigaard, Niels-Ulrik; Wittwer, Matthias; Tonolla, Mauro

*Published in:*  
Standards in Genomic Sciences

*DOI:*  
[10.1186/s40793-018-0317-z](https://doi.org/10.1186/s40793-018-0317-z)

*Publication date:*  
2018

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY](#)

*Citation for published version (APA):*  
Luedin, S. M., Pothier, J. F., Danza, F., Storelli, N., Frigaard, N-U., Wittwer, M., & Tonolla, M. (2018). Complete genome sequence of "Thiodictyon syntrophicum" sp. nov. strain Cad16<sup>T</sup>, a photolithoautotrophic purple sulfur bacterium isolated from the alpine meromictic Lake Cadagno. *Standards in Genomic Sciences*, 13, [14].  
<https://doi.org/10.1186/s40793-018-0317-z>

REPORT

Open Access



# Complete genome sequence of “*Thiodictyon syntrophicum*” sp. nov. strain Cad16<sup>T</sup>, a photolithoautotrophic purple sulfur bacterium isolated from the alpine meromictic Lake Cadagno

Samuel M. Luedin<sup>1,2,3</sup>, Joël F. Pothier<sup>4</sup>, Francesco Danza<sup>1,2</sup>, Nicola Storelli<sup>1,2</sup>, Niels-Ulrik Frigaard<sup>5</sup>, Matthias Wittwer<sup>3</sup> and Mauro Tonolla<sup>1,2\*</sup> 

## Abstract

“*Thiodictyon syntrophicum*” sp. nov. strain Cad16<sup>T</sup> is a photoautotrophic purple sulfur bacterium belonging to the family of *Chromatiaceae* in the class of *Gammaproteobacteria*. The type strain Cad16<sup>T</sup> was isolated from the chemocline of the alpine meromictic Lake Cadagno in Switzerland. Strain Cad16<sup>T</sup> represents a key species within this sulfur-driven bacterial ecosystem with respect to carbon fixation. The 7.74-Mbp genome of strain Cad16<sup>T</sup> has been sequenced and annotated. It encodes 6237 predicted protein sequences and 59 RNA sequences. Phylogenetic comparison based on 16S rRNA revealed that *Thiodictyon elegans* strain DSM 232<sup>T</sup> the most closely related species. Genes involved in sulfur oxidation, central carbon metabolism and transmembrane transport were found. Noteworthy, clusters of genes encoding the photosynthetic machinery and pigment biosynthesis are found on the 0.48 Mb plasmid pTs485. We provide a detailed insight into the Cad16<sup>T</sup> genome and analyze it in the context of the microbial ecosystem of Lake Cadagno.

**Keywords:** Phototrophic sulfur bacteria, *Chromatiaceae*, Sulfur cycling, Meromictic lake, CRISPR, Okenone

## Introduction

PSB belonging to the family of *Chromatiaceae* are generally found at the interface of aerobic and sulfidic-anaerobic zones that are exposed to sunlight such as stagnant, hypertrophic water bodies, littoral zones and bacterial mats [1]. The genus *Thiodictyon* was first described by Winogradsky in 1888 [2] and comprises two type strains, *Thiodictyon elegans* strain DSM 232<sup>T</sup> and *Thiodictyon bacillosum* strain DSM 234<sup>T</sup>. “*Thiodictyon syntrophicum*” sp. nov. strain Cad16<sup>T</sup> is the proposed type strain of the species “*Thiodictyon syntrophicum*” [3] within the family of *Chromatiaceae* of the genus *Thiodictyon* [4]. Cultures of

strain Cad16<sup>T</sup> were isolated from the chemocline of the alpine meromictic Lake Cadagno (Ticino, Switzerland). This lake is characterized by high influx of sulfate, magnesium and calcium in the euxinic monimolimnion which favors the formation of a steep chemocline at 10 to 14 m depth [5, 6]. Within this zone a dense population (up to 10<sup>7</sup> cells per ml in summer) of mainly anaerobic phototrophic sulfur bacteria belonging to the PSB genera *Chromatium*, *Lamprocystis*, *Thiodictyon*, *Thiocystis*, and the GSB *Chlorobium* [7] is responsible for up to 40% of the total CO<sub>2</sub> fixation measured in Lake Cadagno [8]. Strain Cad16<sup>T</sup> has been shown to be highly active in CO<sub>2</sub> fixation both in situ and in vitro [9]. Furthermore, aggregation of strain Cad16<sup>T</sup> with SRB of the genus *Desulfocapsa* has been described [3]. In this publication we describe the first complete genome of strain Cad16<sup>T</sup> providing details especially on CO<sub>2</sub> fixation, sulfur metabolism and on CRISPRs. The sequencing of strain Cad16<sup>T</sup> is part of a larger sequencing project that includes the key species of

\* Correspondence: [mauro.tonolla@supsi.ch](mailto:mauro.tonolla@supsi.ch)

<sup>1</sup>University of Geneva, Sciences III, Department of Botany and Plant Biology, Microbiology Unit, 1211 Geneva, Switzerland

<sup>2</sup>University of Applied Sciences of Southern Switzerland (SUPSI), Department of Environment, Constructions and Design (DACD), Laboratory of Applied Microbiology (LMA), Via Mirasole 22A, 6500 Bellinzona, Switzerland  
Full list of author information is available at the end of the article



the microbial community from the anoxic layers of Lake Cadagno.

## Organism information

### Classification and features

Strain Cad16<sup>T</sup> is Gram-negative, the cells are oval-sphere shaped and 1.4–2.4 µm in diameter, non-motile, vacuolated and contain BChl *a*. Isolate Cad16<sup>T</sup> can grow as single cells, as well as in cell aggregates with up to 100 cells contained in EPS layer (Fig. 1). It was isolated from the chemocline of Lake Cadagno in a depth of 10–14 m where it grows in a non-obligate mutualistic association with sulfur-reducing bacteria of the genus *Desulfocapsa* [10]. Based upon morphology and partial 16S rRNA sequence analysis, the strain Cad16<sup>T</sup> was classified as a member of the genus *Thiodictyon* within the family *Chromatiaceae* before [10]. Figure 2 shows the phylogenetic placement of strain Cad16<sup>T</sup> (complete 16S rRNA sequence) in a 16S rRNA based maximum likelihood phylogenetic tree. The closest relatives of isolate Cad16<sup>T</sup> are *T. bacillosum* DSM 234<sup>T</sup> and *T. elegans* DSM 232<sup>T</sup> with 99% sequence identity (partial 16S rRNA sequences). A comparison of the strain Cad16<sup>T</sup> core genome with other whole genome sequenced PSB confirmed the phylogenetic placement (Additional file 1: Figure S1).

Strain Cad16<sup>T</sup> was anaerobically grown in Pfennigs medium [11], containing per liter: 0.25 g KH<sub>2</sub>PO<sub>4</sub>, 0.34 g NH<sub>4</sub>Cl, 0.5 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.25 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.34 g KCl, 1.5 g NaHCO<sub>3</sub>, 0.5 ml trace element solution SL<sub>10</sub>, and 0.02 mg vitamin B<sub>12</sub> with 2 mM acetate in 100 mL serum bottles with rubber stoppers. The medium was prepared in a 2 l bottle with a N<sub>2</sub>/CO<sub>2</sub> (80%/20%) gas

phase. The medium was then reduced with 0.3 g l<sup>-1</sup> Na<sub>2</sub>S·9H<sub>2</sub>O (1.10 mM final concentration) and adjusted to a pH of 7.2. Cultures were incubated at 20–23 °C under photoheterotrophic conditions with 6 h light/dark photoperiods with a 40-W tungsten bulb placed at a distance of 60 cm from the cultures (ca. 10 µE m<sup>-2</sup> s<sup>-1</sup>).

Different electron donors and carbon substrates were tested under phototautotrophic conditions by Peduzzi et al. [3, 10]. Photolithoautotrophic growth was observed under anoxic conditions with hydrogen sulfide, thiosulfate and elemental sulfur as electron donors. Thereby, elemental sulfur is stored within the periplasma as intermediate oxidation product (Fig. 1). The carbon sources acetate, butyrate, ethanol, formate, fructose, fumarate, glucose, glycerol, lactate, malate, propanol, propionate, pyruvate and succinate were added at 5 mM concentration, respectively. Strain Cad16<sup>T</sup> was observed to assimilate only acetate, pyruvate and fructose in the presence of sulfide and bicarbonate. Strain Cad16<sup>T</sup> was additionally tested for chemolithoautotrophic growth with bicarbonate under a headspace atmosphere containing 5% O<sub>2</sub>, 10% CO<sub>2</sub> and 85% N<sub>2</sub>, in the dark. Growth was observed with 0.02% hydrogen sulfid and 0.07% thiosulfate, or with 0.07% sulfide only, respectively. The pigments responsible for the purple-red color of strain Cad16<sup>T</sup> were analysed spectrometrically in vivo by Peduzzi et al. [3]. Local absorption maxima at 833 nm, 582 nm and 374 nm gave evidence for the presence of BChl *a*, and at 528 nm for the carotenoid okenone, respectively [10].

A further characterization of strain Cad16<sup>T</sup> can be found in Table 1.

A circular representation of the genome sequence and annotation according to the COG criteria is shown in Fig. 3.

## Genome sequencing information

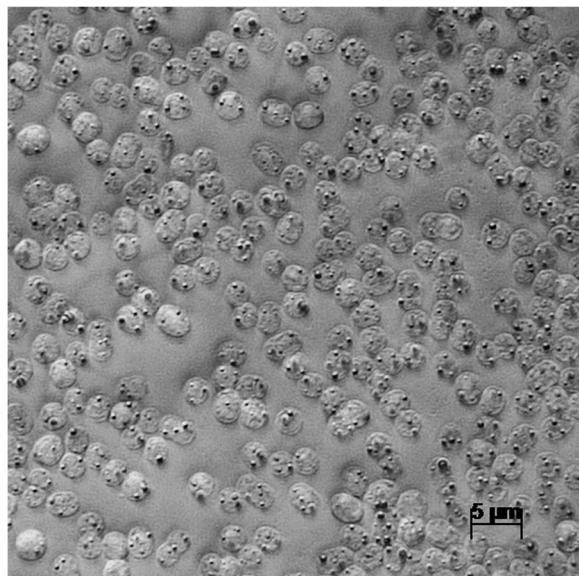
### Genome project history

Sampling was done in August 2001 using a Friedinger-type bottle on Lake Cadagno. Subsequent isolation and cultivation of strain Cad16<sup>T</sup> was done in Pfennig's medium I [11]. gDNA was isolated in November 2014 and sequencing was performed in January 2015. Raw data was assembled in with the SMRTview assembly platform and annotated using the NCBI Prokaryotic Genome Annotation Pipeline. Completeness of the isolate Cad16<sup>T</sup> sequence was verified using the 31 single copy genes of the *Amphora* Net analysis platform [12].

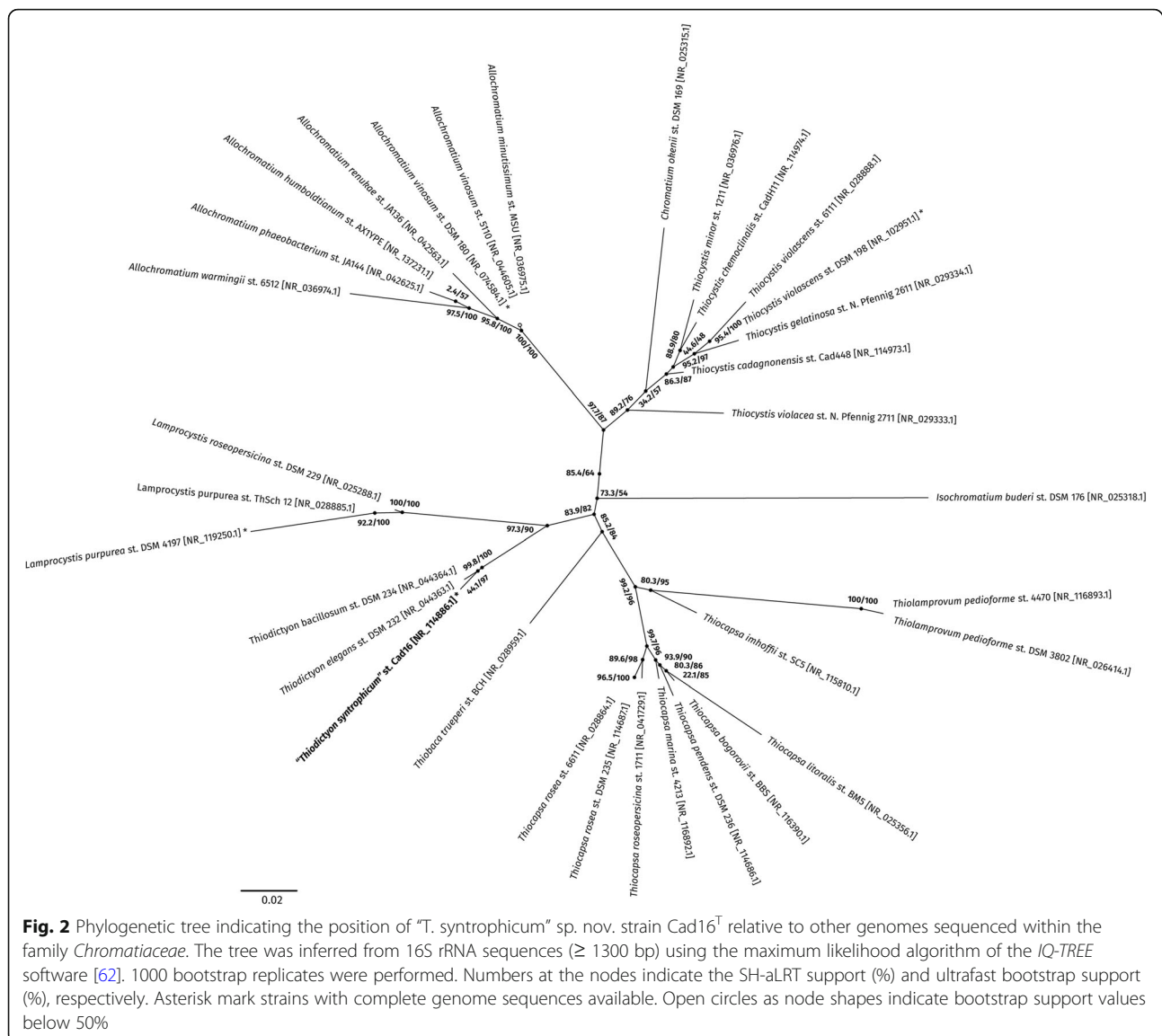
The genome sequence was deposited in GenBank under BioProject [PRJNA354524](#), with the accession numbers CP020370–CP020372. The key elements of the genome studied are listed in Table 2.

### Growth conditions and genomic DNA preparation

Strain Cad16<sup>T</sup> was anaerobically grown in Pfennigs medium [11] Cells were collected by centrifugation for 15 min at 10,600 g. DNA was extracted using phenol/



**Fig. 1** Phase-contrast photomicrograph of “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup>. The elementary sulfur globule inclusions are visible as black dots within the cell



chloroform/isoamylalcohol solution (25:24:1, *v/v*, Sigma, Buchs, Switzerland) following the protocol provided by Pacific Biosciences [13] in combination with phase lock gels (VWR International). gDNA was purified using AMPure beads (Agencourt, Beckman Coulter Life Sciences, Indianapolis, USA) following the E2612 protocol from New England Biolabs [14]. Purity of the DNA was tested using the Qbit UV/VIS absorption reader (Thermo Fisher Scientific, Rheinach, Switzerland).

#### Genome sequencing and assembly

The library construction and genome sequencing was done on the Pacific Biosciences RS II platform at the Functional Genomic Center Zurich, Zurich, Switzerland. A 10 kb SMRTbell library was constructed using the DNA Template Prep Kit 1.0 (Pacific Biosciences, Menlo

Park, USA). SMRTbell template fragments over 10 kb length were used for creating a SMRT bell-Polymerase Complex with P6-C4 chemistry (Pacific Biosciences) according to the manufacturer instructions.

Four SMRT cells v3.0 (Pacific Biosciences) for PacBio RS II chemistry were used for sequencing. Separate sequencing quality reports for all four cells were created through the SMRT portal software.

The SMRT web portal was used for genome assembly with the RS\_HGAP\_Assembly.2 pipeline from the SMRT Analysis 2.3 server. The polished assembly consists of 153 scaffolds with a mean coverage of 175× and a N50 value of 6,849,178. Thereof, three scaffolds were distinctly longer (6.85, 0.50 and 0.43 Mb, respectively) and showed a coverage greater than 200×, whereas mean coverage dropped below a value of 50× for the remaining 150 scaffolds.

**Table 1** Classification and general features of “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup> according to the MIGS recommendations [65]

MIGS ID	Property	Term	Evidence code <sup>a</sup>
	Classification	Domain <i>Bacteria</i>	TAS [3, 59]
		Phylum <i>Proteobacteria</i>	TAS [3]
		Class <i>Gammaproteobacteria</i>	TAS [3]
		Order <i>Chromatiales</i>	TAS [3]
		Family <i>Chromatiaceae</i>	TAS [3]
		Genus <i>Thiodictyon</i>	TAS [2, 60]
		Species “”	TAS [3]
		Strain: Cad16 <sup>T</sup>	TAS [3]
	Gram stain	Negative	TAS [3]
	Cell shape	Coccus	TAS [3]
	Motility	Non-motile	TAS [3]
	Sporulation	No	NAS
	Temperature range	5–25 °C	TAS [3]
	Optimum temperature	20–23	TAS [3]
	pH range; Optimum	6.8–7.5	TAS [3]
	Carbon source	CO <sub>2</sub> , acetate, pyruvate, fructose	TAS [3]
MIGS-6	Habitat	Fresh water, alpine meromictic lake	TAS [3]
MIGS-6.3	Salinity	Not determined	NAS
MIGS-22	Oxygen requirement	Aerotolerant	TAS [3]
MIGS-15	Biotic relationship	Free-living	TAS [3]
MIGS-14	Pathogenicity	Non-pathogen	NAS
MIGS-4	Geographic location	Switzerland, Ticino	TAS [3]
MIGS-5	Sample collection	08.28.2001	TAS [3]
MIGS-4.1	Latitude	46°33′ N	TAS [3]
MIGS-4.2	Longitude	8°43′ E	TAS [3]
MIGS-4.4	Altitude	1923 m	TAS [3]

<sup>a</sup>Evidence codes – *IDA* Inferred from Direct Assay, *TAS* Traceable Author Statement (i.e., a direct report exists in the literature), *NAS* Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [10]

These three scaffolds showed self-similar ends in dot-plot graphs and could be circularized manually.

The genome was manually corrected for SNPs using MiSeq Illumina 300-bp paired-end reads from previous sequencing (unpublished data, N. Storelli, J.F. Pothier, M. Tonolla).

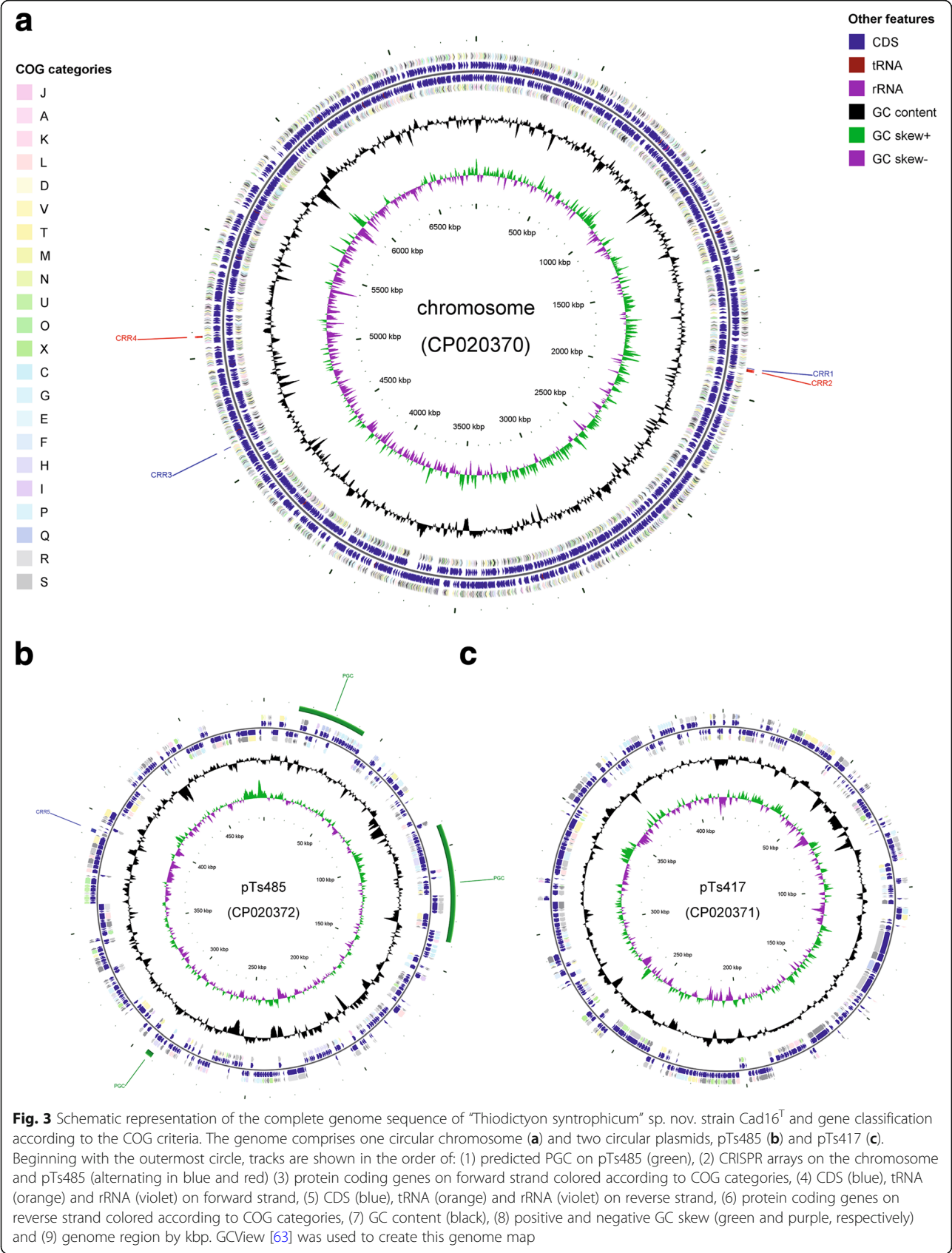
#### Genome annotation

NCBI Prokaryotic Genome Annotation Pipeline (Annotation Software revision 4.1) NCBI Prokaryotic Genome Annotation Pipeline (Annotation Software revision 4.1) was used for gene calling and gene annotation. To identify CRISPR-Cas sequences the CRISPRfinder server was used [15]. The Pfam-A v29 database was used to predict Pfam domains [16]. Transmembrane domains were predicted with the webserver based TMHMM2 program [17] and signal peptides were predicted with SignalP 4.1 server [18].

#### Genome properties

The complete genome of strain Cad16<sup>T</sup> comprises one circular chromosome (6,837,296 bp) and two circular plasmids pTs485 (484,824 bp) and pTs417 (416,864 bp) (Table 3). The average GC content for the chromosome, and plasmids pTs485 and pTs417, is 66.28%, 65.59 and 65.97%, respectively. A total of 6601 coding sequences were predicted. Thereof, 6237 were predicted to encode proteins whereas six rRNA, 49 tRNA and four ncRNA sequences were predicted. A putative function is assigned for 3471 (46.57%) protein encoding genes (Table 4). The classification of genes into COGs functional categories is given in Table 5. The replicons pTs485 and pTs417 could be made circular, have their own origin of replication each, but do not contain any RNA or house-keeping genes. Therefore, to our understanding, both pTs485 and pTs417 fulfill the plasmid definition.





**Table 2** Project information for “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup>

MIGS ID	Property	Term
MIGS 31	Finishing quality	Complete
MIGS-28	Libraries used	SMRT 10 kb (BluePippin size selection)
MIGS 29	Sequencing platforms	PacBio RS II
MIGS 31.2	Fold coverage	200x
MIGS 30	Assemblers	HGAP2
MIGS 32	Gene calling method	GeneMarkS+, software revision: 4.1
	Locus Tag	THSYN
	GeneBank ID	CP020370.1, CP020371.1, CP020372.1
	GeneBank Date of Release	07/12/2017
	GOLD ID	Gp0131589
	BIOPROJECT	PRJNA354524
MIGS 13	Source Material Identifier	NA
	Project relevance	Environmental

### Extended insights from the genome sequence

#### Phototrophy

PSB typically transform light energy into chemical energy with the membrane bound type 2 photochemical reaction center. The chromosome of strain Cad16<sup>T</sup> encodes the core antenna proteins LH1, subunits PufA and PufB (THSYN\_31145 and THSYN\_31140), and the regulatory protein PufQ (THSYN\_31110) upstream to the reaction center genes composed of reaction RC subunits PufL, PufM, and PufC (THSYN\_31125–31,135). Additional two copies of subunits LH2 alpha and beta (THSYN\_31115 and THSYN\_31120), respectively, are encoded further downstream, as well as pairwise in two other clusters (THSYN\_30995/31005/31030/31040 and THSYN\_31000/31005/31010/31035/31045), similar as described for the PSB *Allochrochromatium vinosum* DSM 180<sup>T</sup> [19]. The photosynthetic reaction center H subunit PuhA (THSYN\_31405) and PucC (THSYN\_31410) are clustered upstream with genes encoding RC-LH1 auxiliary proteins (THSYN\_31390–31,400). Furthermore, a homologous Hi-PIP (THSYN\_25970) is found in strain Cad16<sup>T</sup>. It may function as the main electron donor to the photosynthetic reaction center similar as in *A. vinosum* [20].

The absorption spectrum of strain Cad16<sup>T</sup> shows strong absorption peaks at 374 nm, 582 nm and 833 nm which are characteristic for BChl *a* [10]. The genes for the complete enzymatic pathway from protoporphyrin

to chlorophyllide, and further to BChl *a* (THSYN\_31090–31,105, THSYN\_31375, THSYN\_31385, THSYN\_31415–31,445, THSYN\_31555, THSYN\_32265–32,270), are clustered on pTs485. BChl *a* formation is thereby catalyzed by an anaerobic type of the Mg-protoporphyrin IX monomethyl ester oxidative cyclase (ChlE) (THSYN\_31385) and a light independent proto-chlorophyllide reductase complex (ChlLNB) (THSYN\_31420–31,430) in strain Cad16<sup>T</sup>.

**Table 4** Genome statistics for the “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup> genome

Attribute	Value	% of Total
Genome size (bp)	7,738,984	100.00
DNA coding (bp)	6,663,511	86.10
DNA G + C (bp)	5,124,386	66.22
DNA scaffolds	3	100.00
Total genes	6601	100.00
Protein coding genes	6237	94.49
RNA genes	59	0.89
rRNA genes	6	0.09
tRNA genes	49	0.74
ncRNA genes	4	0.06
Pseudo genes	305	4.62
Genes in internal clusters	NA	NA
Genes with function prediction	2737	41.46
Genes assigned to COGs	3157	47.83
Genes with Pfam domains	4675	70.82
Genes with signal peptides	436	6.61
Genes with transmembrane helices	1185	17.95
CRISPR repeats	5	–

NA not applicable

**Table 3** Summary of genome of “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup>: one circular chromosome and two circular plasmids

Label	Size (Mb)	Topology	INSDC identifier	RefSeq ID
Chromosome	6.84	Circular	CP020370	NA
pTs417	0.42	Circular	CP020371	NA
pTs485	0.49	Circular	CP020372	NA

**Table 5** Number of genes associated with general COG functional categories of the genome of “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup>. The percentage is set relative to the total number of protein coding genes in the genome

Code	Value	% age	Description
<u>J</u>	210	3.37	Translation, ribosomal structure and biogenesis
<u>A</u>	1	0.02	RNA processing and modification
<u>K</u>	144	2.31	Transcription
<u>L</u>	276	4.43	Replication, recombination and repair
<u>B</u>	0	0.00	Chromatin structure and dynamics
<u>D</u>	42	0.67	Cell cycle control, Cell division, chromosome partitioning
<u>Y</u>	0	0.00	Nuclear structure
<u>V</u>	186	2.98	Defense mechanisms
<u>T</u>	297	4.76	Signal transduction mechanisms
<u>M</u>	274	4.39	Cell wall/membrane biogenesis
<u>N</u>	10	0.16	Cell motility
<u>Z</u>	0	0.00	Cytoskeleton
<u>W</u>	0	0.00	Extracellular Structures
<u>U</u>	62	0.99	Intracellular trafficking and secretion
<u>O</u>	226	3.62	Posttranslational modification, protein turnover, chaperones
<u>X</u>	97	1.56	Energy production and conversion
<u>C</u>	245	3.93	Energy production and conversion
<u>G</u>	122	1.96	Carbohydrate transport and metabolism
<u>E</u>	167	2.68	Amino acid transport and metabolism
<u>F</u>	49	0.79	Nucleotide transport and metabolism
<u>H</u>	135	2.16	Coenzyme transport and metabolism
<u>I</u>	93	1.49	Lipid transport and metabolism
<u>P</u>	184	2.95	Inorganic ion transport and metabolism
<u>Q</u>	29	0.46	Secondary metabolites biosynthesis, transport and catabolism
<u>R</u>	308	4.94	General function prediction only
<u>S</u>	1522	24.40	Function unknown
No COG	1543	24.74	Not in COGs
Multi COG	320	5.13	Multiple COG assignments

Strain Cad16<sup>T</sup> produces okenone as its sole carotenoid [10] and Crt proteins involved in carotenoid biosynthesis are found on pTs485. The complete synthesis of this keto-carotenoid is mediated through two novel types of carotenoid ketolases, the C-4/4' ketolase CruO (THSYN\_31065) and the oxygen dependent CruS bifunctional desaturase (THSYN\_31070) [21]. The characteristic  $\chi$ -ring of okenone is introduced through the key enzymes CrtY and CrtU (THSYN\_31055 and THSYN\_31050) [21, 22].

Remarkably, most of the proteins involved in photosynthesis are encoded on plasmid pTs485, forming a PGC (Fig. 3) [23]. The highly modular character of the *pufLM* and *pufC* genes of  $\alpha$ ,  $\beta$  and  $\gamma$ -proteobacteria has been demonstrated previously [24, 25]. To our knowledge, this is the first description of a PGC being localized on a plasmid in a PSB species. Interestingly, the gene cluster is similarly organized as in the  $\gamma$ -proteobacterium

*Congregibacter litoralis* strain KT71<sup>T</sup> and as in members from the  $\alpha$ -proteobacteria families *Rhodobacteraceae* and *Rhodospirillaceae*, respectively.

#### Sulfur metabolism

For the photoautotrophic process of CO<sub>2</sub> assimilation in PSB, electrons derived from the oxidation of reduced sulfur compounds, are transferred to electron carriers NAD(P)<sup>+</sup> and ferredoxin through light energy. During photolithoautotrophic growth under anaerobic conditions, strain Cad16<sup>T</sup> uses electrons from the oxidation of sulfide, thiosulfate and elemental sulfur as reducing equivalents [3]. Strain Cad16<sup>T</sup> can use thiosulfate as an electron source during phototrophic growth [3]. No homologous genes for the thiosulfate oxidizing multi-enzyme complex SoxAX, could be found in the strain Cad16<sup>T</sup> genome. However, *soxB* (THSYN\_26690) and clustered genes



encoding SoxYZ (THSYN\_09005–09010) that binds thiosulfate were identified in the genome. Remarkably, this gene combination is found in several genome sequenced *Ectothiorhodospiraceae*. In contrast to the PSB *A. vinosum* DSM 180<sup>T</sup> [26], no homologous sequence for the tetrathionate-forming thiosulfate dehydrogenase TsdA was found. However, a c4 cytochrome type TsdB homolog (THSYN\_17090) was identified. Due to this unusual combination of genes involved in thiosulfate oxidation, further studies are needed to elucidate the thiosulfate oxidation pathways in strain Cad16<sup>T</sup>.

Initial sulfide and thiosulfate oxidation is immediately followed SGB formation in strain Cad16<sup>T</sup> (Fig. 1). In strain Cad16<sup>T</sup> the SGB structure is mediated through envelope SGP homologues to SgpA and SgpB (THSYN\_20250 and THSYN\_05960) from “*Thioflaviccoccus mobilis*” and *Thiocystis violascens*, respectively. The sequence of SgpC (THSYN\_11025) shows homology to *Marichromatium* species SgpC/CV3. Predicted signal peptides suggest export of for all three SGP proteins into the periplasm in Cad16<sup>T</sup>, as proposed for *A. vinosum* DSM 180<sup>T</sup> [27].

Moreover, the genome of strain Cad16<sup>T</sup> encodes the membrane-bound sulfide: quinone oxidoreductases SqrD (THSYN\_04215) and SqrF (THSYN\_09305). These are possibly involved in the oxidation of sulfide in the periplasm.

The mode of sulfur transport across the inner membrane is not known for PSBs [28]. Organic persulfides such as glutathione or glutathione amide persulfide are proposed as possible candidates. In a next step, the rhodanese-like protein Rhd transfers the sulfur from the persulfide-carrier to the TusA protein in the cytoplasm. The further oxidation steps from sulfur to sulfite are typically mediated through the reverse acting *dsr* genes in PSB [29]. The strain Cad16<sup>T</sup> genes in the *dsr* cluster (THSYN\_22480, THSYN\_22490–22,545) are arranged in a highly conserved organization similar to *A. vinosum* DSM 180<sup>T</sup>, only missing *dsrS* that is non-essential for sulfur oxidation [30]. The DsrEFH complex mediates persulfate transfer from TusA onto DsrC. The persulfurated form of DsrC is then substrate for the cytoplasmic reverse-acting dissimilatory sulfite reductase DsrAB that catalyzes the formation of sulfite. Finally, DsrMKJOP complex reduces DsrC [30].

The genome harbors three additional sulfur relay proteins similar to DsrC (THSYN\_09485, THSYN\_18820 and THSYN\_22565) that could function as TusA homologues. In *A. vinosum* DSM 180<sup>T</sup> DsrC is able to bind DNA upstream the *dsr* cluster [31].

In strain Cad16<sup>T</sup>, *soeABC* (THSYN\_16370–16,380) encode the sulfur-iron molybdoprotein complex that further oxidizes sulfite to sulfate on the cytoplasmic site of the membrane [32]. Alternatively, strain Cad16<sup>T</sup> oxidizes sulfite via APS by APS-reductase AprBA (THSYN\_16395 and THSYN\_16400) and ATP sulfurylase Sat

(THSYN\_16390), as in other PSB [33, 34]. Thereby, the membrane-bound QmoABHdrCB-complex [35] (THSYN\_16425–6440) possibly functions as an electron acceptor for the AprAB reductase complex since no *aprM* homolog was found in the strain Cad16<sup>T</sup> sequence. For the extra-cytoplasmic export of the final oxidation product sulfate, a SulP sulfate permease (THSYN\_14085) homolog to *A. vinosum* DSM 180<sup>T</sup> is encoded in the strain Cad16<sup>T</sup> sequence.

Hydrogen uptake and consumption has been shown to be linked to sulfur metabolism in *Thiocapsa roseopersicina* BBS [36, 37]. Thereby, electrons from hydrogen oxidation in the periplasm by the hyn-type hydrogenase HydSL could be transferred via the Isp membrane complex to the disulfide bound to DsrC. In *A. vinosum* DSM 180<sup>T</sup>, transcription of *isp1* and *isp2* encoding the Isp hydrogenase subunits is upregulated during growth on sulfide [38]. The Isp complex is composed of two subunits, Isp1 and Isp2, that contain similar catalytic domains as DsrM and DsrK, respectively. Similarly, homologous Isp1 and Isp2 proteins (THSYN\_28105 and THSYN\_28100) may link sulfur to hydrogen metabolisms in strain Cad16<sup>T</sup>. In accordance, an increase in the sulfide concentration was observed while SGB were consumed by strain Cad16<sup>T</sup> during incubation in the dark (unpublished results, F. Danza).

Additionally, other [NiFe]-hydrogenases of the Hox and Hup type (THSYN\_22655, THSYN\_22660 and THSYN\_28115) are found in the sequence that could mediate light-dependent H<sub>2</sub> evolution as proposed for *T. roseopersicina* [39, 40].

The Cad16<sup>T</sup> genome also harbors *cys* genes (THSYN\_05020–05035) that are probably involved in sulfate assimilation under sulfur-limiting conditions. Furthermore, the genome also encompasses genes encoding the CydDC (THSYN\_18930 and THSYN\_18935) ATP-driven cysteine transport proteins [41].

### Autotrophic growth

In PSB, CO<sub>2</sub> fixation is essentially achieved through the reductive pentose phosphate also known as the CBB cycle. In accordance, the strain Cad16<sup>T</sup> genome harbors the complete CBB enzymatic pathway. On the chromosome, the dimeric RuBis-CO form II (THSYN\_13250) clusters with RuBis-CO activation protein subunits CbbR, CbbQ and CbbO, (THSYN\_13245, THSYN\_13255 and THSYN\_13285). Interestingly, small and large RuBis-CO subunits form I (THSYN\_29475 and THSYN\_29480) cluster together with carboxysome shell and auxiliary proteins on plasmid pTs417 (THSYN\_29485–29,520 and THSYN\_29530–29,535). The carboxysome may allow efficient photoassimilation across varying CO<sub>2</sub> concentrations as proposed for *A. vinosum* DSM 180<sup>T</sup> [42]. Previous studies showed

different expression regulation for RuBis-CO type I and type II genes from Cad16<sup>T</sup> suggesting that only the type II is involved in the process of CO<sub>2</sub> fixation [8]. Interestingly, the plasmid pTs485 also harbors a RuBis-CO-like protein form III gene (THSYN\_31160) upstream the PGC.

The missing sedoheptulose-1,7-bisphosphatase SBP is possibly bypassed by via the fructose-1,6-bisphosphatase (THSYN\_25630). The genes *gltA* citrate synthase (THSYN\_12620), *fumA* fumarate hydratase (THSYN\_24360) and *sucCD* succinyl-CoA ligase (THSYN\_00880 and THSYN\_00885) that are essential for the TCA cycle, and isocitrate lyase (THSYN\_16275) and malate synthase (THSYN\_15655) that are essential for the glyoxylate cycle, respectively, are identified in the strain Cad16<sup>T</sup> sequence. Recently a proteomic study about the capacity of Cad16<sup>T</sup> to fix CO<sub>2</sub> in the dark suggested the presence of a particular archaeal DC/HB cycle [42]. However, no further genes coding for this DC/HB cycle were found. Also a complete set of genes coding for polyhydroxyalkanoic acid synthase PhaC (THSYN\_06910) and poly-(3-hydroxybutyrate) depolymerase PhaE (THSYN\_06905) are found in the strain Cad16<sup>T</sup> genome.

Strain Cad16<sup>T</sup> additionally encodes genes necessary for glycogen polymerisation. The glucose 1-phosphate adenylyltransferase GlgC (THSYN\_00810), the glycogen synthase GlgA (THSYN\_11615) and the 1,4- $\alpha$ -glucan branching enzyme GlgB (THSYN\_00805) allow the synthesis of glycogen.

Interestingly, strain Cad16<sup>T</sup> also has the potential to produce the storage compound cyanophycin normally found in *cyanobacteria* [43], since the two subunits of the enzyme cyanophycin synthetase (THSYN\_26990 and THSYN\_26995) are found.

Togther, these finding provide genetic evidence for the high carbon fixation potential of strain Cad16<sup>T</sup> in the dark [8, 44].

Anaerobic Fe(II)-oxidation was described for other *Thiodictyon* strains [45, 46] and evidence of cryptic in situ iron cycling has been demonstrated recently [47]. In accordance with these findings, we found *cbh3* type terminal cytochrome C oxidases (THSYN\_06760–08775) possibly involved in Fe(II) driven carbon fixation in strain Cad16<sup>T</sup> genome.

Strain Cad16<sup>T</sup> grows chemoautotrophically under microaerobic conditions (5% O<sub>2</sub>) with sulfide, thiosulfate, or sulfide only [3], as also observed in other PSB in vitro studies with *Lamprocystis purpurea* [10, 48], *Thiocystis violacea* and *A. vinosum* [49]. In situ, strain Cad16<sup>T</sup> is possibly exposed to low concentration of oxygen produced by oxygenic microbiota at the mixolimnion-chemocline interface [8]. Accordingly, we observe genes encoding *sod*-type superoxide dismutases (THSYN\_20405 and THSYN\_22720), as well as *fir* and *fur*-type

transcriptional regulators involved in peroxide stress response. In situ, strain Cad16<sup>T</sup> is possibly exposed to oxygen produced by oxygenic microbiota at the mixolimnion-chemocline interface [8].

### Nitrogen metabolism

Furthermore, with the genes encoding NifB (THSYN\_03975), NifD (THSYN\_08880), NifH (THSYN\_08885), NifK (THSYN\_08875), NifT (THSYN\_08870) NifW, NifZ and NifM (THSYN\_10720, THSYN\_10725 and THSYN\_10730), NifX (THSYN\_21435) and NifL (THSYN\_24590) strain Cad16<sup>T</sup> could possibly fix nitrogen. Genes encoding the multisubunit urease UreDEFG (THSYN\_03745, 03750, 03760 and 03765) and the urea transporter UrtABCDE (THSYN\_07940–07955, 03760, 07975) indicate the possible utilisation of urea.

### Transmembrane transport proteins

Several membrane transport genes were found in the strain Cad16<sup>T</sup> genome, including protein secretion system Type II, genes encoding the TAT pathway and several TRAP transporter genes, as well as genes encoding Ton-Tol type and ABC-type transporter complexes. Additionally, a complete TSS4 pilus machinery is encoded in six clusters dispersed on the strain Cad16<sup>T</sup> chromosome. Notably, also structural components of TSS6 secretion system are found in two clusters on the chromosome (THSYN\_11395–11,410) and on pTs485 (THSYN\_32540–THSYN\_32580). Two effector proteins of the VrgG family were identified. THSYN\_15360 belongs to the *vgr\_GE* type Rhs family proteins similar sequences found in  $\beta$ -proteobacterial family of the *Burkholderiaceae* whereas THSYN\_32425 is conserved in  $\gamma$ -proteobacteria and contains a type IV Rhs element. Togther, the secretion machinery allows strain Cad16<sup>T</sup> to interact within the highly populated chemocline with up to 10<sup>7</sup> bacterial cells per milliliter. The secretion and uptake mechanism may also play a key role in the cell-to-cell contact with *Desulfocapsa thiozymogenes*.

### Buoyancy regulation and chemotaxis

Strain Cad16<sup>T</sup> can possibly regulate buoyancy by gas vesicles that are formed with the encoded structural gas vesicle proteins. Whereas GvpA proteins forms the vesicle core (THSYN\_11790, THSYN\_11825, THSYN\_15290, THSYN\_18705 and THSYN\_31215), GvpFL (THSYN\_11800 and THSYN\_18685), GvpK (THSYN\_11785) and GvpN (THSYN\_11815 and THSYN\_18695) further stabilize the structure. Proteins homologues to the transcriptional regulatory factors GvrA (THSYN\_11850) and GvrC (THSYN\_11830) from the enterobacterium *Serratia* sp. ATCC 39006 are also found in Cad16<sup>T</sup>.

The diurnal and seasonal behavior of vacuolated *Chromatiaceae* has been described for different lakes [50, 51].

In strain Cad16<sup>T</sup> a diguanylate cyclase (THSYN\_19835) is found upstream the circadian clock genes *kaiCBB* (THSYN\_19820–19,830). These genes act together [52] and may synchronize optimal flotation within the chemocline.

### CRISPR-Cas systems

Bacterial CRISPR-Cas systems provide a mechanism against bacteriophage infection and plasmid transformation [53]. A CRISPR locus is composed out of a 300–500 bp leader sequence, spacer sequences (21–72 bp), complementary to foreign DNA, and direct repeats (DRs, 24–40 bp) flanking them [53–55]. Adjacent *cas* genes encode protein that are co-transcribed with the CRISPR locus and interfere with invading DNA guided by the specific spacers [56, 57].

Five CRISPR repeat regions (CRR1–CRR5) were identified in the genome of strain Cad16<sup>T</sup>, four being located on the chromosome and one on the plasmid pTs485 (Fig. 3). The number of DRs ranges from 19 (CRR4) to 146 (CRR2) as seen in Table 6.

BLASTn analysis of the CRISPR DRs using the CRISPRfinder platform revealed similarities in CRR1, CRR2 and CRR4 to sequences of “*T. mobilis*” 8321 (57 hits, 2 mismatches) and “*Thioalkalivibrio sulfidophilus*” HL-EbGr7 (63 hits, 3 mismatches). The DRs found in CRR3 are similar to the ones in *Halothiobacillus neapolitanus* c2 (31 hits, 4 mismatches), whereas the DRs in CRR5 are similar to the ones found in *Vibrio alginolyticus* NBRC 15630 (1 hit, 5 mismatches).

Furthermore, three CRISPR-Cas loci were identified in the strain Cad16<sup>T</sup> sequence, containing *cas3* genes that are characteristic for type I CRISPR-Cas systems [58]. A complete CRISPR-Cas loci (THSYN\_08045–08070) is located 201 bp upstream of CRR2 and is assigned to subtype I-U, containing the signature protein (THSYN\_08055) of the *GSU0054* family (TIGR02165 and a *cas3*, THSYN\_08070) with C-terminal HD domain (TIGR01596) [58]. Another CRISPR array (THSYN\_19240–19,290) is located 182 bp upstream of CRR3 and is classified as subtype I-C due to the *cas8c* gene and the lack of a *cas6* sequence. Additionally, an incomplete CRISPR-Cas locus (CRR5) is identified on plasmid pTs485, encoding for Cas2, Cas1,

(THSYN\_19240–19,245, THSYN\_19265, THSYN\_19275, THSYN\_19285 and THSYN\_19,290).

### Conclusions

We report on the first complete genome sequence of “*Thiodictyon syntrophicum*” sp. nov. strain Cad16<sup>T</sup> and the metabolic versatility of this environmentally relevant organism. The observed carbon fixation potential can be explained by the highly developed photosynthesis machinery that is coupled to the sulfur and carbon metabolism. Within the changing conditions in the chemocline, strain Cad16<sup>T</sup> is able to optimally use light, different organic and inorganic carbon compounds, reduced sulfur, nitrogen and oxygen. The two 0.4 Mb plasmids found in Cad16<sup>T</sup> are unique for known PSB species and we report structural similarity to sequences from  $\alpha$ - and  $\gamma$ -proteobacterial phototrophs. The availability of the complete genome sequence of strain Cad16<sup>T</sup> will facilitate further studies that elucidate its role as key species of the chemocline and the tight association with the *Desulfocapsa* sp. and the interaction with different PSB and GSB species present in the anoxic part of Lake Cadagno. Due to the limited molecular data on other *Thiodictyon* strains and no reference strains available, no (digital) DNA-DNA hybridization experiments could be performed. However, the result from phylogenetic analyses on 16S rRNA sequence level, comparative genomic analyses as well as morphological and physiological differences (see above) indicate a novel species within the genus *Thiodictyon*.

The described isolate is therefore proposed as “*Thiodictyon syntrophicum*” sp. nov. strain Cad16<sup>T</sup>, a novel species within the genus *Thiodictyon*.

A formal description of the proposed novel species follows below:

#### Description of “*Thiodictyon syntrophicum*” sp. nov.

“*Thiodictyon syntrophicum*” (syn.tro'phi.cum. Gr. pref. Syn, together with; Gr. adj. *Trophikos*, nursing, tending or feeding; N.L. neut. Adj. *syntrophicum*, syntrophic).

Gram-negative, cells are oval-round shaped and 1.4–2.4  $\mu$ m in diameter, non-motile, vacuolated and contain BChl *a* and okeneone. Growth as single cells, as well as in aggregates with up to 100 cells in a EPS layer.

**Table 6** CRISPR-Cas loci detected in “*T. syntrophicum*” sp. nov. strain Cad16<sup>T</sup> genome

Localization	Name	CRISPR start	CRISPR end	CRISPR length (bp)	DR consensus	DR length	No. of spacers	CRISPR-Cas loci <sup>a</sup>
chromosome	CRR1CRR1	1,879,131	1,881,639	2508	GCTTCAATGAGGCCGCGCGAATTCGCCGCGGAAAC	36	34	type I-U
	CRR2CRR2	1,883,646	1,894,325	10,679	GCTTCAATGAGGCCGCGCGAATTCGCCGCGGAAAC	36	146	
	CRR3CRR3	4,626,522	4,629,249	2727	GCATCGCCCGGCCAATTGGCCGGGCGCGGATTGAAAC	37	37	
	CRR4CRR4	5,078,034	5,085,199	7165	GTTTCGCGGCGAATTCGCCGCGCCTCATTGAAGC	36	98	–
pTs485	CRR5CRR5	391,741	393,104	1363	GTAGCGTACTCCGAGCCGCAAGGCTATTGAAAC	35	19	–

<sup>a</sup> CRISPR-Cas classification according to Makarova et al. [58]



Assimilation of elemental sulfur in intracellular sulfur globules. Grow photoautotrophically in Pfennig's minimal medium with a doubling time of 121 h at 20–23 °C, a pH of 6.8–7.2, at 1 mM sulfide and a photoperiod of 12 h dark/ 12 h light. Dense cultures show a milky purple-red and milky color. Carbon assimilation via Calvin cycle. Following carbon substrates were utilized at a concentration of 5 mM: acetate, fructose and pyruvate. No growth was observed with 5 mM butyrate, ethanol, formate, fumarate, glucose, glycerol, lactate, malate, propanol, propionate and succinate, respectively. Chemolithoautotrophic growth was observed with 5% Oxygen and 0.02% hydrogen sulfide and 0.07% thiosulfate, or with 0.07% sulfide only, respectively.

The type strain Cad16<sup>T</sup> (=JCM 15483<sup>T</sup> =KCTC5955<sup>T</sup>) was isolated from a sulfidic chemocline in the alpine Lake Cadagno in Switzerland. The genome size of the type strain is 6.84 Mb (chromosome), contains two plasmids, pTs485 (0.49 Mb) and pTs417 (0.42 Mb) and the G + C content of the genome is 66.22%. The 16S rRNA gene sequence of strain Cad16<sup>T</sup> is deposited under the GenBank/EMBL/DBJ accession number AJ511274. The complete genome sequence of the type strain Cad16<sup>T</sup> is deposited under the GenBank ID CP020370, CP020371 and CP020372. The type strain has been deposited both at the Japan Collection of Microorganisms (JCM 15483<sup>T</sup>) and at the Korean Collection for Type Cultures (KCTC 5955<sup>T</sup>).

## Additional file

**Additional file 1: Figure S1.** Phylogenetic placement of “*T. syntrophicum*” strain Cad16<sup>T</sup> within the other 12 Chromatiaceae species with a publicly available whole genome sequences. Additionally, the closely related phylogenetic lineages *Nitrosococcus*, *Rheinheimera* and *Arsukibacterium* are also included. Strain Cad16<sup>T</sup> is most closely related to *L. purpurea* DSM 4197. The maximum likelihood tree was inferred from 100 concatenated single-copy orthologues sequences [61] and a total of 1000 bootstrap replicates were performed. Numbers at the nodes indicate the SH-aLRT support (%) and ultrafast bootstrap support (%). OrthoMCL [64], was used to define a set of orthologues proteins between these 23 species. Hundred single-copy orthologues were randomly chosen and aligned with MUSCLE [66]. The best-fit phylogenetic model and subsequent consensus tree computation, based on maximum-likelihood and 1000 bootstrap iterations, was performed with the IQ-TREE software [62]. Nodes with both, 100% SH-aLRT and ultrafast bootstrap support, are indicated with filled black circle symbols for convenience. (TIF 57220 kb)

## Abbreviations

ABC: ATP-binding cassette; APS: adenosine-5'-phosphosulfate; BChl *a*: bacteriochlorophyll *a*; Cas: CRISPR associated; CBB: Calvin–Benson–Bassham cycle; CRISPRs: clustered regularly interspaced short palindromic repeats; DC/ HB: dicarboxylate/4-hydroxybutyrate cycle; DR: direct repeat; *dsr*: dissimilatory sulfite reductase; EPS: extracellular polymeric substances; GSB: green sulfur bacteria; HiPIP: high-potential iron-sulfur protein; PGC: photosynthesis gene cluster; PSB: purple sulfur bacteria; RC: reaction center; RubisCO: ribulose-1,5-bisphosphate carboxylase/oxygenase; SGB: sulfur globule; SMRT: single molecule real-time; SRB: sulfur reducing bacteria; T4P: type IV pilus; T6SS: type VI protein secretion system; TAT: twin-arginine translocation; TCA: tricarboxylic acid cycle; TRAP: tripartite ATP-independent periplasmic

## Acknowledgements

We kindly thank Nicole Liechti, Corinne Oechslin, Pierre H.H. Schneeberger and Christian Beuret. We also thank the assistance of Andrea Patrignani from the Functional Genomics Center Zurich for performing the PacBio RS II sequencing.

## Authors' contributions

SML and FD drafted the manuscript, carried out cultivation, DNA extraction and purification and microscopy. SML, JFP and MW participated in the genome assembly, correction and annotation. SML, JFP, FD, NS, MW, NUF and MT discussed and analyzed the data and revised the manuscript. MT, JFP and MW conceived of and supervised the study. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>University of Geneva, Sciences III, Department of Botany and Plant Biology, Microbiology Unit, 1211 Geneva, Switzerland. <sup>2</sup>University of Applied Sciences of Southern Switzerland (SUPSI), Department of Environment, Constructions and Design (DADC), Laboratory of Applied Microbiology (LMA), Via Mirasole 22A, 6500 Bellinzona, Switzerland. <sup>3</sup>Federal Office for Civil Protection, Spiez Laboratory, Biology Division, Spiez, Switzerland. <sup>4</sup>Zurich University of Applied Sciences (ZHAW), Institute of Natural Resource Sciences, Environmental Genomics and System Biology Research Group, Wädenswil, Switzerland. <sup>5</sup>University of Copenhagen, Department of Biology, Helsingør, Denmark.

Received: 24 July 2017 Accepted: 24 April 2018

Published online: 09 May 2018

## References

- Imhoff JF. The Family Chromatiaceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. *The Prokaryotes* [Internet]. Springer Berlin Heidelberg; 2014 [cited 2017 Apr 24]. p. 151–78. Available from: [http://link.springer.com/referenceworkentry/10.1007/978-3-642-38922-1\\_295](http://link.springer.com/referenceworkentry/10.1007/978-3-642-38922-1_295).
- Winogradsky S. Beiträge zur Morphologie und Physiologie der Bakterien. Heft 1. Zur Morphologie und Physiologie der Schwefelbakterien. Leipzig: Felix; 1888.
- Peduzzi S, Storelli N, Welsh A, Peduzzi R, Hahn D, Perret X, et al. *Candidatus "Thiodictyon syntrophicum"*, sp. nov., a new purple sulfur bacterium isolated from the chemocline of Lake Cadagno forming aggregates and specific associations with *Desulfocapsa* sp. *Syst Appl Microbiol*. 2012;35:139–44.
- Imhoff JF. *Chromatiales* Ord. Nov. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, Boone DR, De Vos P, et al., editors. *Bergey's manual® Syst. Bacteriol. Vol. two Proteobacteria part B Gammaproteobacteria*. Boston: Springer US; 2005. p. 1–59. Available from: [http://dx.doi.org/10.1007/0-387-28022-7\\_1](http://dx.doi.org/10.1007/0-387-28022-7_1)
- Del Don C, Hanselmann KW, Peduzzi R, Bachofen R. The meromictic alpine Lake Cadagno: orographical and biogeochemical description. *Aquat Sci*. 2001;63:70–90.
- Tonolla M, Storelli N, Danza F, Ravasi D, Peduzzi S, Posth NR, et al. Lake Cadagno: Microbial Life in Crenogenic Meromixis. In: Gulati RD, Zareev ES, Degermendzhi AG, editors. *Ecol Meromictic Lakes*. Cham: Springer International Publishing; 2017 [cited 2017 Jun 16]. p. 155–86. Available from: [http://link.springer.com/10.1007/978-3-319-49143-1\\_7](http://link.springer.com/10.1007/978-3-319-49143-1_7)
- Tonolla M, Peduzzi R, Hahn D. Long-term population dynamics of phototrophic sulfur bacteria in the chemocline of Lake Cadagno, Switzerland. *Appl Environ Microbiol*. 2005;71:3544–50.
- Camacho A, Erez J, Chicote A, Florin M, Squires MM, Lehmann C, et al. Microbial microstratification, inorganic carbon photoassimilation and dark carbon fixation at the chemocline of the meromictic Lake Cadagno (Switzerland) and its relevance to the food web. *Aquat Sci*. 2001;63:91–106.
- Storelli N, Peduzzi S, Saad MM, Frigaard N-U, Perret X, Tonolla M. CO<sub>2</sub> assimilation in the chemocline of Lake Cadagno is dominated by a few types of phototrophic purple sulfur bacteria. *FEMS Microbiol Ecol*. 2013;384:421–32.
- Peduzzi S, Tonolla M, Hahn D. Isolation and characterization of aggregate-forming sulfate-reducing and purple sulfur bacteria from the chemocline of meromictic Lake Cadagno, Switzerland. *FEMS Microbiol Ecol*. 2003;45:29–37.
- Eichler B, Pfennig N. A new purple sulfur bacterium from stratified freshwater lakes, *Amoebobacter purpureus* sp. nov. *Arch Microbiol*. 1988; 149:395–400.

12. AmphoraNet: The webserver implementation of the AMPHORA2 metagenomic workflow suite [Internet]. [cited 2016 Jun 5]. Available from: <http://www.sciencedirect.com/science/article/pii/S0378111913014091>
13. SharedProtocol-Extracting-DNA-usinig-Phenol-Chloroform.pdf [Internet]. [cited 2017 Jan 5]. Available from: <http://www.pacb.com/wp-content/uploads/2015/09/SharedProtocol-Extracting-DNA-usinig-Phenol-Chloroform.pdf>
14. Agencourt AMPure XP Bead Clean-up - NEBNext Microbiome DNA Enrichment Kit (E2612) | NEB [Internet]. [cited 2017 Jan 5]. Available from: <https://www.neb.com/protocols/2013/04/18/agencourt-ampure-xp-bead-clean-up-e2612>
15. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 2007;35:W52–7.
16. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Misty J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 2016;44:D279–85.
17. TMHMM Server, v. 2.0 [Internet]. [cited 2017 Apr 6]. Available from: <http://www.cbs.dtu.dk/services/TMHMM/>
18. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8:785–6.
19. Corson GE, Nagashima KVP, Matsuura K, Sakuragi Y, Wettasinghe R, Qin H, et al. Genes encoding light-harvesting and reaction center proteins from *Chromatium vinosum*. *Photosynth Res.* 1999;59:39–52.
20. Verméglio A, Li J, Schoepp-Cothenet B, Pratt N, Knaff DB. The role of high-potential iron protein and cytochrome *c*<sub>8</sub> as alternative electron donors to the reaction center of *Chromatium vinosum*<sup>†</sup>. *Biochemistry (Mosc).* 2002;41:8868–75.
21. Vogl K, Bryant DA. Elucidation of the biosynthetic pathway for Okenone in *Thiodictyon* sp. CAD16 leads to the discovery of two novel carotene ketolases. *J Biol Chem.* 2011;286:38521–32.
22. Vogl K, Bryant DA. Biosynthesis of the biomarker okenone:  $\chi$ -ring formation. *Geobiology.* 2012;10:205–15.
23. Bauer CE, Buggy JJ, Yang ZM, Marrs BL. The superoperon organization of genes for pigment biosynthesis and reaction center proteins is a conserved feature in Rhodobacter capsulatus: analysis of overlapping bchB and puhA transcripts. *Mol Gen Genet.* 1991;228:433–44.
24. Nagashima KVP, Verméglio A, Fusada N, Nagashima S, Shimada K, Inoue K. Exchange and complementation of genes coding for photosynthetic reaction center core subunits among purple bacteria. *J Mol Evol.* 2014;79:52–62.
25. Igarashi N, Harada J, Nagashima S, Matsuura K, Shimada K, Nagashima KVP. Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple Bacteria. *J Mol Evol.* 52:333–41.
26. Denkmann K, Grein F, Zigann R, Siemen A, Bergmann J, van Helmont S, et al. Thiosulfate dehydrogenase: a widespread unusual acidophilic c-type cytochrome. *Environ Microbiol.* 2012;14:2673–88.
27. Pattaragulanit K, Brune DC, Trüper HG, Dahl C. Molecular genetic evidence for extracytoplasmic localization of sulfur globules in *Chromatium vinosum*. *Arch Microbiol.* 1998;169:434–44.
28. Frigaard N-U, Dahl C. Sulfur Metabolism in Phototrophic Sulfur Bacteria. In: Poole RK, editor. *Adv. Microb. Physiol.* [Internet]. Academic Press; 2008 [cited 2016 Mar 16]. p. 103–200. Available from: <http://www.sciencedirect.com/science/article/pii/S0065291108000027>
29. Pott AS, Dahl C. Sirohaem sulfite reductase and other proteins encoded by genes at the *dsr* locus of *Chromatium vinosum* are involved in the oxidation of intracellular sulfur. *Microbiology.* 1998;144:1881–94.
30. Grein F, Pereira IAC, Dahl C. Biochemical characterization of individual components of the *Allochrochromatium vinosum* DsrMKJOP transmembrane complex aids understanding of complex function in vivo. *J Bacteriol.* 2010;192:6369–77.
31. Grimm F, Dobler N, Dahl C. Regulation of *dsr* genes encoding proteins responsible for the oxidation of stored sulfur in *Allochrochromatium vinosum*. *Microbiology.* 2010;156:764–73.
32. Dahl C, Franz B, Hensen D, Kesselheim A, Zigann R. Sulfite oxidation in the purple sulfur bacterium *Allochrochromatium vinosum*: identification of SoxABC as a major player and relevance of SoxYZ in the process. *Microbiology.* 2013;159:2626–38.
33. Hipp WM, Pott AS, Thum-Schmitz N, Faath I, Dahl C, Trüper HG. Towards the phylogeny of APS reductases and sirohaem sulfite reductases in sulfate-reducing and sulfur-oxidizing prokaryotes. *Microbiology.* 1997;143:2891–902.
34. Parey K, Demmer U, Warkentin E, Wymen A, Ermler U, Dahl C. Structural, biochemical and genetic characterization of dissimilatory ATP Sulfurylase from *Allochrochromatium vinosum*. *PLoS One.* 2013;8:e74707.
35. Meyer B, Kuever J. Homology modeling of dissimilatory APS reductases (AprBA) of sulfur-oxidizing and sulfate-reducing prokaryotes. *PLoS One.* 2008;3:e1514.
36. Tengölics R, Mészáros L, Györi E, Doffkay Z, Kovács KL, Rákhely G. Connection between the membrane electron transport system and Hyn hydrogenase in the purple sulfur bacterium, *Thiocapsa roseopersicina* BBS. *Biochim Biophys Acta BBA - Bioenerg.* 1837;2014:1691–8.
37. Laurinavichene T. The effect of sulfur compounds on H<sub>2</sub> evolution/consumption reactions, mediated by various hydrogenases, in the purple sulfur bacterium, *Thiocapsa roseopersicina* - Springer [Internet]. [cited 2016 Jun 20]. Available from: <http://link.springer.com/article/10.1007%2Fs00203-007-0260-7>
38. Weissgerber T, Sylvester M, Kröninger L, Dahl C. A comparative quantitative proteomic study identifies new proteins relevant for sulfur oxidation in the purple sulfur bacterium *Allochrochromatium vinosum*. *Appl Environ Microbiol.* 2014;80:2279–92.
39. Rákhely G, Laurinavichene TV, Tsygankov AA, Kovács KL. The role of Hox hydrogenase in the H<sub>2</sub> metabolism of *Thiocapsa roseopersicina*. *Biochim Biophys Acta BBA - Bioenerg.* 2007;1767:671–6.
40. Rákhely G, Kovács ÁT, Maróti G, Fodor BD, Csanádi G, Latinovics D, et al. Cyanobacterial-type, Heteropentameric, NAD<sup>+</sup>-reducing NiFe hydrogenase in the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina*. *Appl Environ Microbiol.* 2004;70:722–8.
41. Pittman MS, Corker H, Wu G, Binet MB, Moir AJG, Poole RK. Cysteine is exported from the *Escherichia coli* cytoplasm by CydDC, an ATP-binding cassette-type transporter required for cytochrome assembly. *J Biol Chem.* 2002;277:49841–9.
42. Weissgerber T, Zigann R, Bruce D, Chang Y-J, Detter JC, Han C, et al. Complete genome sequence of *Allochrochromatium vinosum* DSM 180(T). *Stand Genomic Sci.* 2011;5:311–30.
43. Simon RD. Cyanophycin granules from the blue-green alga *Anabaena cylindrica*: a reserve material consisting of copolymers of aspartic acid and arginine. *Proc Natl Acad Sci U S A.* 1971;68:265–7.
44. Storelli N, Saad MM, Frigaard N-U, Perret X, Tonolla M. Proteomic analysis of the purple sulfur bacterium Candidatus “Thiodictyon syntrophicum” strain Cad16T isolated from Lake Cadagno. *EuPA Open Proteomics* [Internet]. [cited 2013 Nov 28]; Available from: <http://www.sciencedirect.com/science/article/pii/S2212968513000172>
45. Ehrenreich A, Widdel F. Anaerobic oxidation of ferrous iron by purple bacteria, a new type of phototrophic metabolism. *Appl Environ Microbiol.* 1994;60:4517–26.
46. Croal LR, Johnson CM, Beard BL, Newman DK. Iron isotope fractionation by Fe(II)-oxidizing photoautotrophic bacteria. *Geochim Cosmochim Acta.* 2004;68:1227–42.
47. Berg J. Intensive cryptic microbial iron cycling in the low iron water column of the meromictic Lake Cadagno - Berg - 2016 - Environmental Microbiology - Wiley Online Library [Internet]. [cited 2016 Oct 23]. Available from: <http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13587/full>
48. Overmann J, Pfennig N. Continuous chemotrophic growth and respiration of *Chromatiaceae* species at low oxygen concentrations. *Arch Microbiol.* 1992;158:59–67.
49. Kampf C, Pfennig N. Capacity of chromatiaceae for chemotrophic growth. Specific respiration rates of *Thiocystis violacea* and *Chromatium vinosum*. *Arch Microbiol.* 1980;127:125–35.
50. Overmann J, Pfennig N. Bouyancy regulation and aggregate formation in *Amoebobacter purpureus* from Mahoney Lake. *FEMS Microbiol Lett.* 1992;101:67–79.
51. Egli K, Wiggli M, Fritz M, Klug J, Gerss J, Bachofen R. Spatial and temporal dynamics of a plume of phototrophic microorganisms in a meromictic alpine lake using turbidity as a measure of cell density. *Aquat Microb Ecol.* 2004;35:105–13.
52. York A. Structural biology: the tick-tock of circadian clocks. *Nat Rev Microbiol.* 2017;15:256–7.
53. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* 2007;315:1709–12.
54. Mojica FJM, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of *Archaea*, *Bacteria* and mitochondria. *Mol Microbiol.* 2000;36:244–6.
55. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology.* 2005;151:2551–61.



56. Al-Attar S, Westra ER, van der Oost J, SJJ B. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol Chem.* 2011;392:277–89.
57. Marraffini LA. CRISPR-Cas immunity in prokaryotes. *Nature.* 2015;526:55–61.
58. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol.* 2015;13:722–36.
59. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A.* 1990;87:4576–9.
60. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. *Int J Syst Evol Microbiol.* 1980;30:225–420.
61. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Nat Genet.* 2000;25:25–9.
62. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol.* 2015;32:268–74.
63. Grin I, Linke D. GCVIEW: the genomic context viewer for protein homology searches. *Nucleic Acids Res.* 2011;39 (suppl):W353–W356.
64. Chen F. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* 2006;34(90001):D363–D368.
65. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, dePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, Hermjakob H, Hertz-Fowler C, Hugenholtz P, Joint I, Kagan L, Kane M, Kennedy J, Kowalchuk G, Kottmann R, Kolker E, Kravitz S, Kyrpides N, Leebens-Mack J, Lewis JE, Li K, Lister AL, Lord P, Maltsev N, Markowitz V, Martiny J, Methe B, Mizrahi I, Moxon R, Nelson K, Parkhill J, Proctor L, White O, Sansone S-A, Spiers A, Stevens R, Swift P, Taylor C, Tateno Y, Tett A, Turner S, Ussery D, Vaughan B, Ward N, Whetzel T, Gil IS, Wilson G, Wipat A. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol.* 2008;26(5):541–547.
66. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–1797.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

